FORCAST Spectroscopy Pipeline Tutorial

This tutorial is designed to help guide users through the steps of processing FORCAST spectroscopy data, from raw data files to the final, calibrated data product. The information in this tutorial can also be found in more detail FORCAST User's Manual. Additional detail about some of the data reduction modules can also be found in Herter et al. (2013; 10.1086/674144).

Retrieve the Data

For this sample reduction, we'll use FORCAST data from the planetary nebula NGC 7009 (the Saturn Nebula). These data were obtained using the FOR_G227 grism, which covers the 17.6 - 26.6 μm wavelength range. These observations were taken for project 05_0063 (PI: Guzman-Ramirez) in 2017. To retrieve the data from the SOFIA Science Archive at IRSA, enter the following criteria in the main SOFIA search page (https://irsa.ipac.caltech.edu/applications/sofia):

  Spatial Constraints → All-Sky
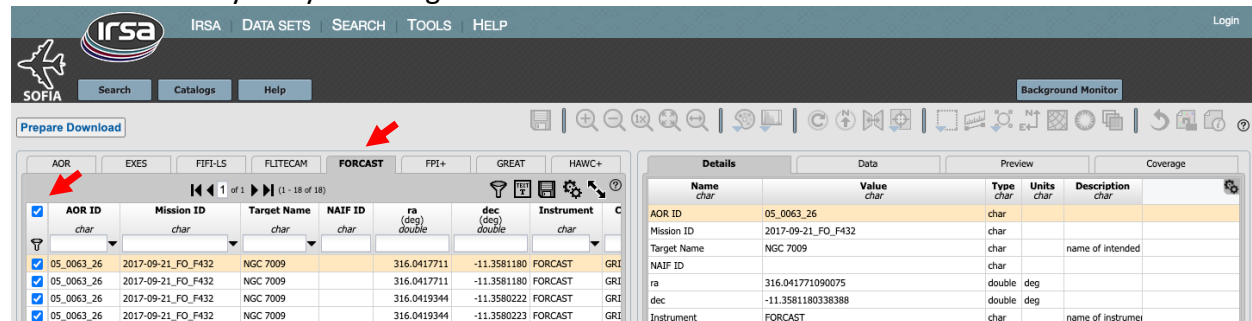 Observation Constraints → AOR ID → 05_0063_26
  Data Product Constraints → Select Level 1 only (de-select Level 3 and Level 4)
or follow this link:

https://irsa.ipac.caltech.edu/applications/sofia/?api=search&spatialConstraints=allsky&execute=true&processingLevel=LEVEL_1&aorId=05_0063_26

On the search results page, click on the FORCAST tab. On the FORCAST tab, select all 18 data files, either individually or by selecting the box next to the AOR ID column header.

Click on Prepare Download to open a download dialog box, and then select Prepare Download within the dialog box. After a few moments, you will be asked to save a zip file to disk (It is ~15MB in size). Save this file wherever you wish, unzip it, and navigate the subfolders to locate the 18 FORCAST FITS files. You may either leave these files in the directory or move them to another location of your choice; the folder containing the 18 FITS files will be referred to as the 'input' directory for the remainder of the tutorial.
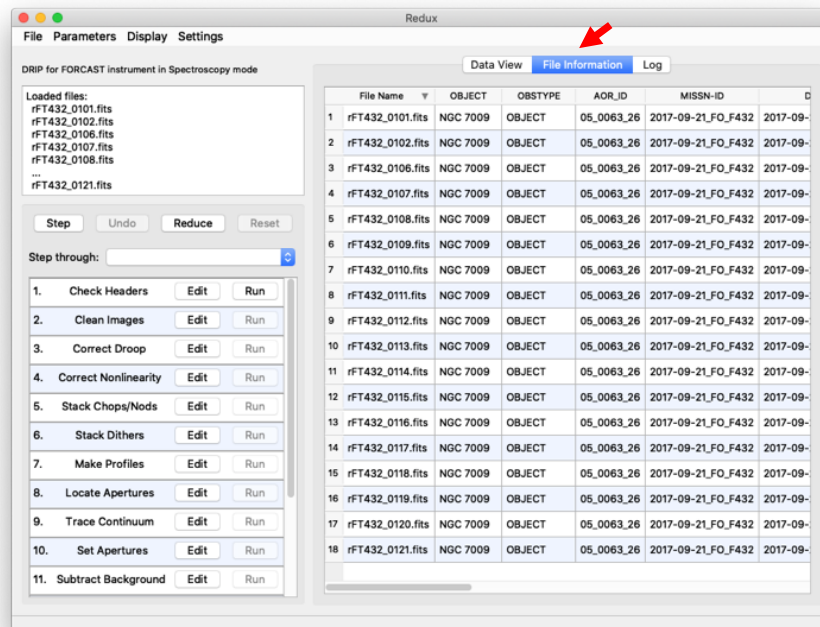
Download and Install the Pipeline

Instructions on how to download and install the *sofia_redux* pipeline package can be found at https://github.com/SOFIA-USRA/sofia_redux. If you need assistance, you can contact us through GitHub or the SOFIA Helpdesk (sofia_help@sofia.usra.edu).

Start the GUI

The pipeline can be executed in either Automatic Mode (on the command line) or in Manual Mode (through a GUI). In this tutorial, we will focus on running the pipeline through the GUI. To begin, launch the GUI by opening a terminal and typing:
  > redux

In the Redux GUI, load the NGC 7009 data by selecting File → Open New Reduction, navigate to the input directory where you saved the data, select all 18 FITS files by clicking on the first file and shift-clicking on the last file, and then click Open. After clicking Open, the software reads the FITS headers and extracts relevant information, such as the instrument name for determining which pipeline the software will call (note that 'DRIP for FORCAST instrument in Spectroscopy mode' is now displayed above the list of loaded files). Some additional FITS header information is shown on the File Information on the right side of the GUI.

When the pipeline runs, it is possible to automatically display the intermediate data using SAO DS9 if it is set up properly (see http://ds9.si.edu/ for download and installation instructions; the ds9 executable must be available in the PATH environment variable for the pyds9 interface to be able to find and control it). Alternatively, the automatic display can be disabled (each pipeline module can optionally produce FITS file output in case a different FITS viewer is preferred). Options for displaying images in DS9 are found under Display Settings on the Data View tab.
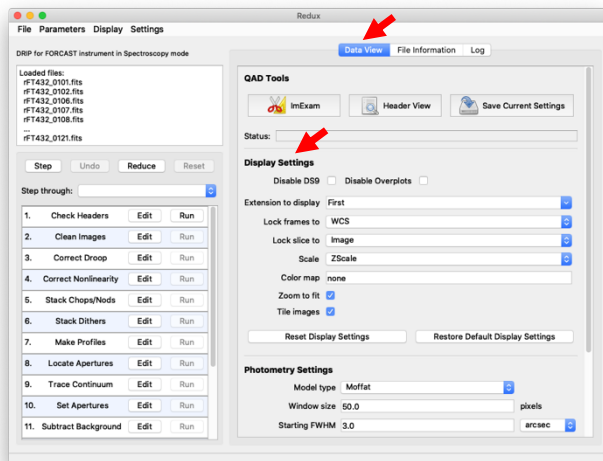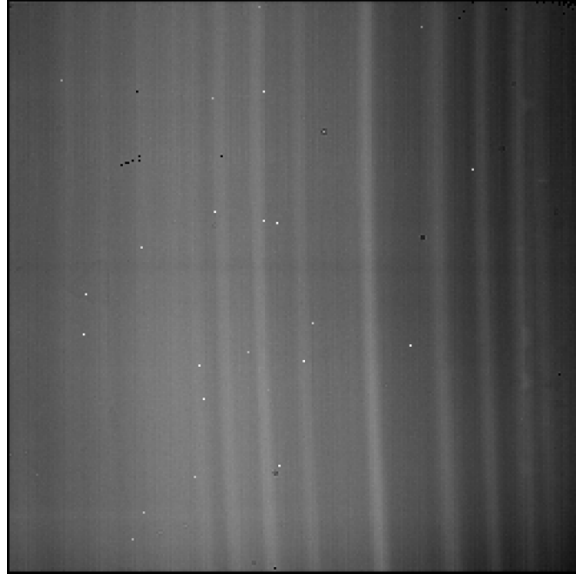
Run the Pipeline

There are 16 modules in the FORCAST spectroscopy pipeline, shown in the bottom-left panel of the GUI. Each pipeline module can be configured by clicking the 'Edit' button for that module. The entire pipeline can be executed by clicking the 'Reduce' button. Alternatively, the modules can be run individually by clicking on the 'Run' button for that module, or in a small group by selecting an end point in the 'Step Through' dropdown menu, then clicking 'Step.' In this tutorial, we'll run the modules one by one and examine their output.

1. Check Headers: This module examines the fits headers to make sure that they have the keywords required by the pipeline.

Options: Click on the Edit button next to Check Header module. A popup window for Edit Parameters for Check Headers will open up. In the popup window if the 'Abort reduction for invalid headers' option is not selected, the pipeline may fail in unexpected ways (as always: garbage in, garbage out!).
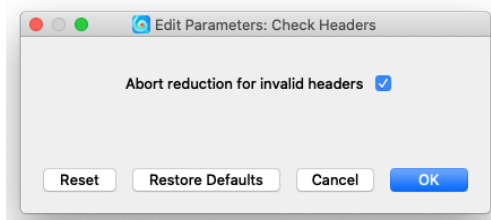

Output: This module creates no output; it allows the data set to pass through to the next modules in the reduction. For comparison with the output of the next module, we show the first plane (each raw data file in C2N mode has four frames, two chops and two nods) of file rFT432_0101.fits before any pipeline modules have been applied.

First plane of file rFT432_0108.fits

2. Clean Images: This module will apply a bad pixel mask to the data (as NaN values), and search for (and correct) a FORCAST-specific issue in which the data are shifted by 16 pixels.

Options: The pipeline automatically loads the bad pixel map appropriate for the data that will be processed (based on the FITS header information). In most cases, users will want to use the default values in this module.
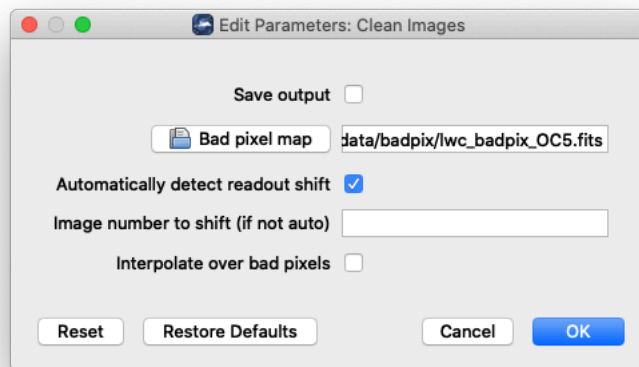

Output: A careful comparison of this image with the prior image shows that the 'hot' pixels are now blanked, along with another elongated region near the center of the array (examples marked with red arrows).
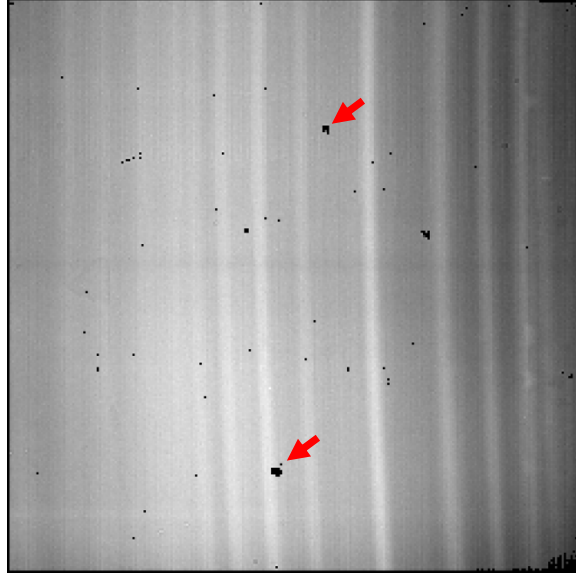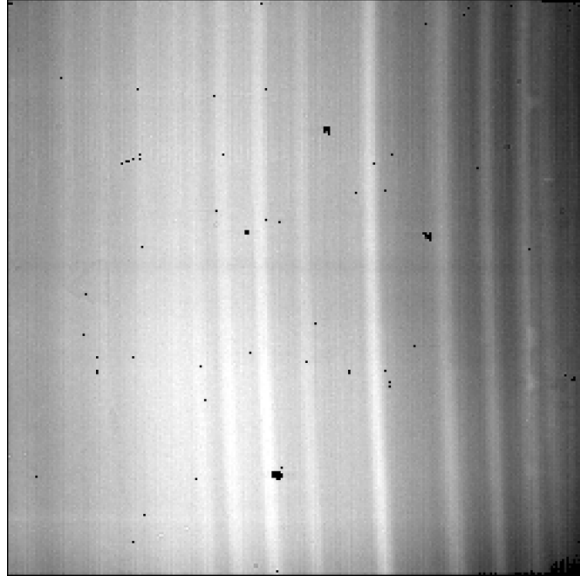
First plane of F0432_FO_GRI_05006326_FORG227_CLN_0108.fits.

3. and 4. Correct Droop and Correct Nonlinearity: These modules correct a few issues specific to the FORCAST arrays and readout electronics (see the User's Manual for more details; see also Fig. 4 of the Herter et al. (2013) for additional examples of how these data issues are corrected).

Options: The default values used in these two modules have been determined by scientists at the SOFIA Science center. You will probably not want to change them without a very good reason.
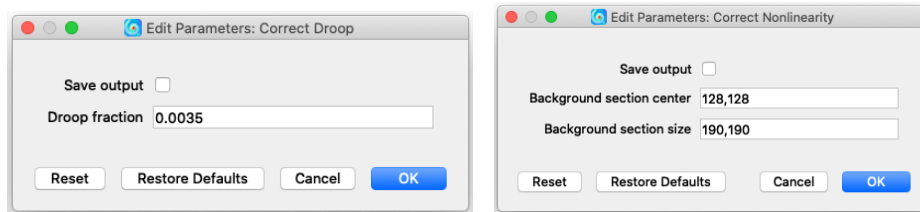

Output: The effects of these two pipeline modules are relatively subtle for this data set.

First plane of F0432_FO_GRI_05006326_FORG227_LNZ_0108.fits.

5. Stack Chops/Nods: This module combines the 4 chop/nod frames into a single image. It also by default removes the 'jailbar' features seen in the prior modules.

Options: As with the earlier modules, the default options are usually best for most data sets.
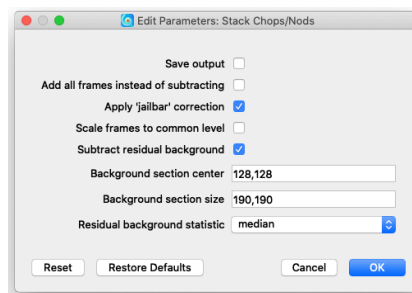

Output: In the image below, the y-axis is along the slit (the spatial direction), and the x-axis is the wavelength axis. The bright region in the image shown by the red arrow indicates that a spectral line has been detected.
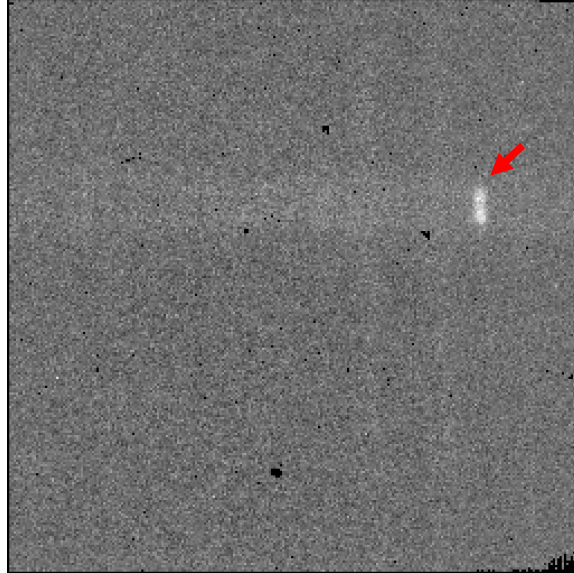
F0432_FO_GRI_05006326_FORG227_STK_0108.fits.

6. Stack Dithers: If the source is very faint, then we can perform a dither stacking module. Because we can see the source in the previous image, this module can be skipped.

Options: The default option is to skip dither stacking. If it is skipped, the other parameters are ignored.
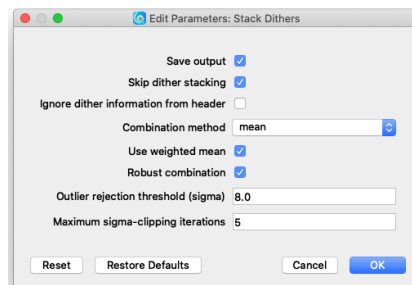

Output: Because the stacking was skipped, this module has no new output.

7. Make Profiles: This module creates a rectified image in which the data are resampled onto a regular spatial-spectral grid. It also creates and displays the smoothed spatial profile of the source across the slit.

Options: The relevant wavelength calibration and slit correction files (if they exist) are loaded by default, but can be specified by the user. In this example, no slit correction file is used. The 'Row fit order' is used to calculate the spatial profile. Order 3 typically works well, but a higher or lower order may yield better results. The 'Subtract median background' will optionally subtract a background from the profile. We will leave this box checked and examine the results.
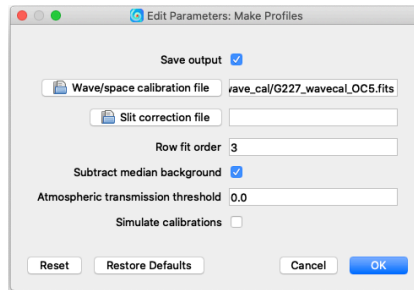
Output: In the image below, the y-axis is along the slit (the spatial direction), and the x-axis is the wavelength axis.
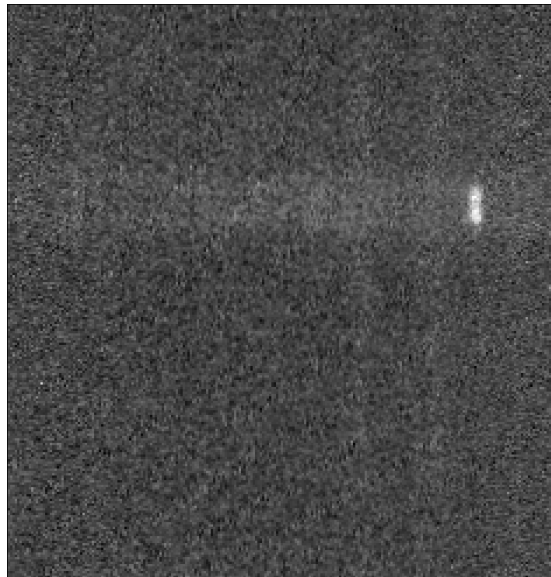
F0432_FO_GRI_05006326_FORG227_RIM_0108.fits.

In addition to the FITS file output, this module (and several of the next modules) also displays results in a pop-up basic spectral viewer. The spectral viewer displays the relevant data (shown in the title) in a single column format and the user can use the scroll bar to view the various data or resize the spectral viewer as shown below:
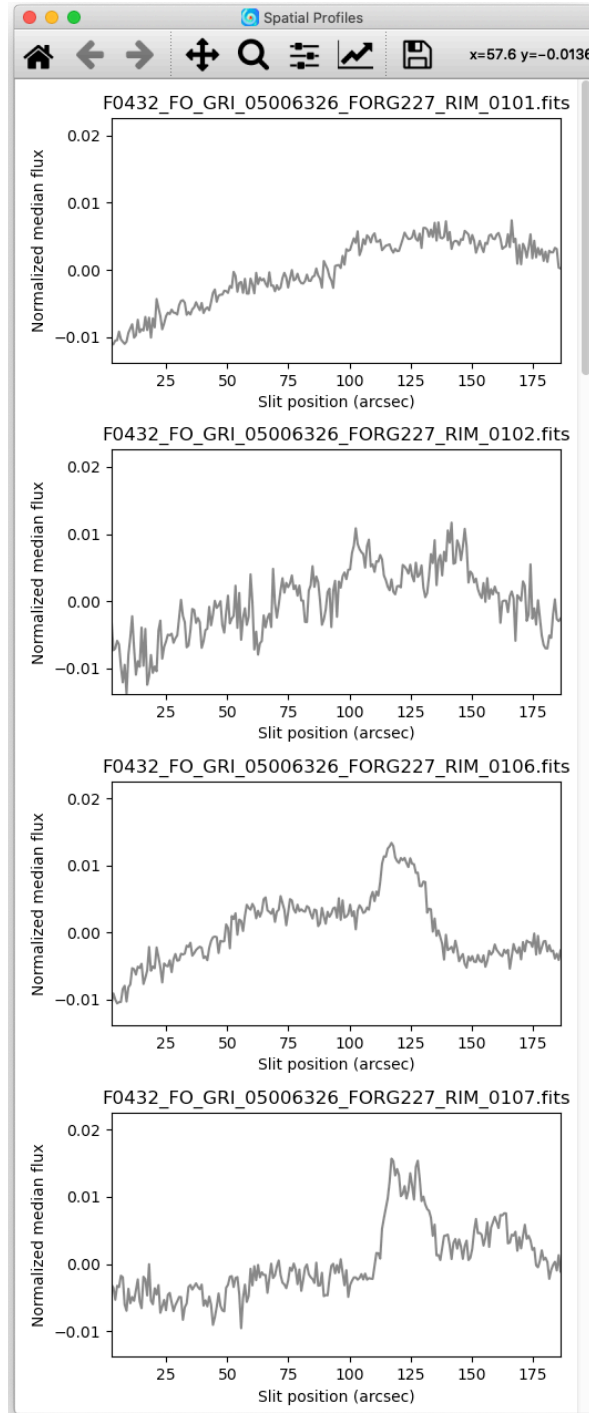
As you can see in this Spatial Profiles pop-up window, the first two profiles are significantly different from the other 16 profiles. The QA comments for the processed data, currently available in the headers of the processed data for this project at IRSA, indicate that for the first two data files (0101 and 0102) the source was off the slit. Therefore, these two files should be excluded from the final data product.

In addition, we can see that several of the profiles have values that are consistently below zero. As noted in the pipeline user's manual, if the profile goes below zero this module should be re-run with the 'Subtract median background' unchecked.

** To exclude the two bad files and re-run module 7, go to the 'File' menu on the top left of the redux GUI and select 'Remove Files' from the pull-down menu, which will open a window with the list of files that are part of the current reduction. Select files rFT432_0101.fits and rFT432_0102. Click OK to exclude these files from processing. Next, click on the 'Edit' button for module 7 (Make Profile Module) and uncheck the 'Subtract median background' box. To run the pipeline through the end of the 'Make Profiles' module, select 'Make Profiles' in the 'Step through' drop-down menu, then click on the 'Step' button. The two bad files will now be excluded from further processing, and we can continue processing the good data.***

8. Locate Apertures: This module identifies the location of the source along the slit.

Options: The default 'auto' option attempts to find the location of the strongest Gaussian peak in the spatial profile. If this does not work for all spectra, users can also specify a specific location along the slit (using 'fix to input') or simply use the central position (using 'fix to center'). It is also possible to fit multiple sources. Since we only have one positive image of our source, we will use one aperture. If the data were obtained by chopping on-slit (e.g., nod-match-chop mode), then we would use more than one aperture.

For this example, we'll try the 'auto' method first, with no Aperture position entered.
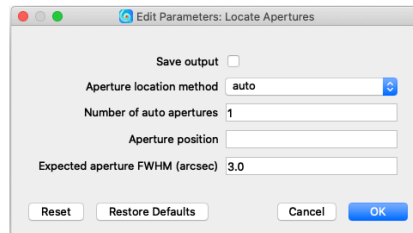

Output: The image below shows the results when the pipeline automatically determines the aperture position. For many of the individual spectra the automatic detection worked well, but for others (e.g., file 0115) it did not because this source has no continuum and shows emission from just one spectral line.
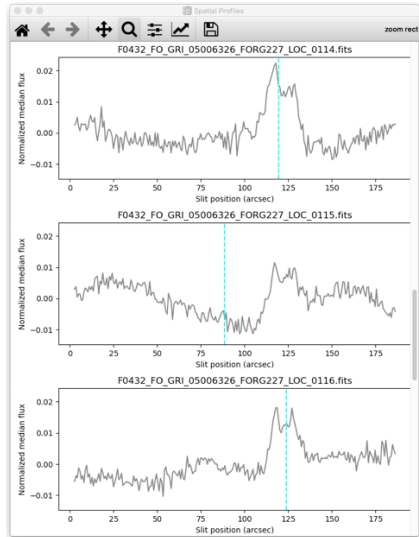
Based on the fits that worked well, we can see that the source is centered at a slit position of ~124 arcsec. We can undo just the Locate Apertures module (by the clicking 'Undo' button), and then enter an initial guess of 124 arcsec in the Aperture position field (this constraint helps the algorithm identify the local peak). We then re-run this module, which results in the following output:

The position of the source is now properly identified for each observation.
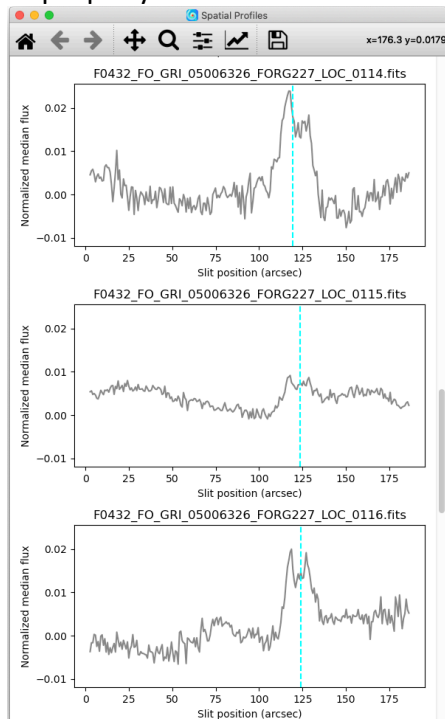

8. Trace Continuum: This module identifies the location of the spectrum across the array, to account for the slit curvature across the array.

Options: The main options are either to have the pipeline determine the location of the spectrum automatically or to fix it to the aperture position. If fixed to the aperture position, a default slit curvature will be used. We will use the 'fit to continuum' method for a more accurate determination of the curvature.
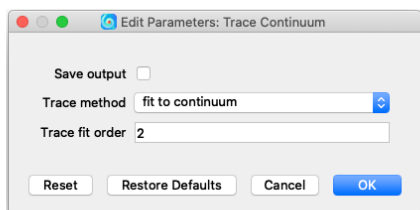

Output: The image below shows the output of this module. The green 'x' positions mark the position of the source at each wavelength, and the red curve is the polynomial fit to the 'x' positions.
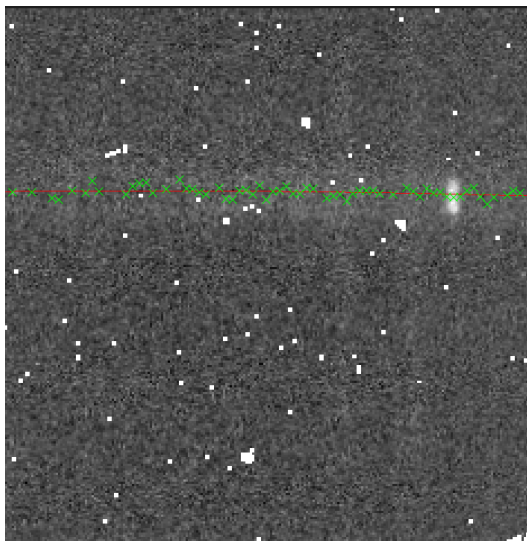

F0432_FO_GRI_05006326_FORG227_TRC_0108.fits overlaid with ds9 region file
F0432_FO_GRI_05006326_FORG227_TRC_0108.reg.

10. Set Apertures: This module identifies the region that will be extracted from each spatial profile to create the final spectrum.

Options: By default, the pipeline automatically determines the aperture and PSF radius to extract by analyzing the spatial profile and the aperture position defined above. It's also possible to override the automatic functionality by specifying these values. For this tutorial, we can use the automatic option.
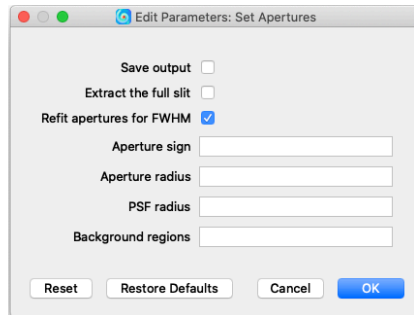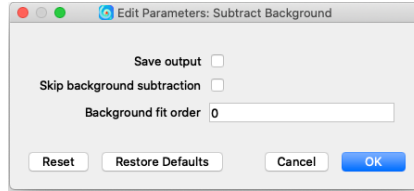
Output: The results of the aperture fitting are shown in the pop-up window, shown below. The cyan line shows the aperture center, the green lines show the edges of the extraction aperture, and the dark blue line shows the region as the flux goes to zero (PSF radius). The red lines indicate the background regions.
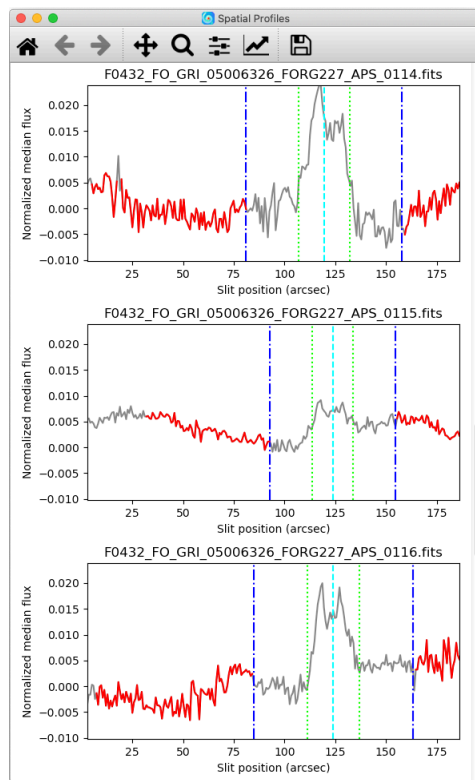

11. Subtract Background: In this module, the residual background is removed using the regions defined in the previous module.

Options: By default, the background is removed, but this module can also be skipped. In our example, the background levels are already close to zero so this module does not make a big difference in the final product. Nevertheless, we will leave the background subtracting on.
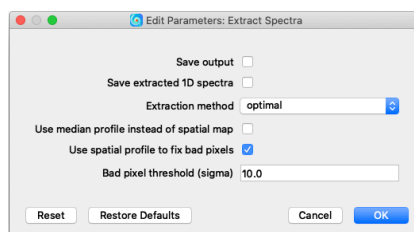
Output: Here is the output from this module, which is the background-subtracted 2-d image. The cyan line shows the aperture center, the green lines show the edges of the extraction aperture, and the dark blue lines show the region as the flux goes to zero (PSF radius), and the red lines mark the background regions. As with the other 2-d images shown, the y-axis is along the slit (the spatial direction), and the x-axis is the wavelength axis.
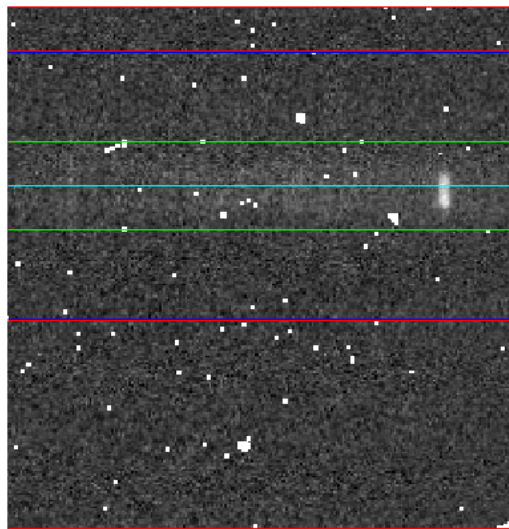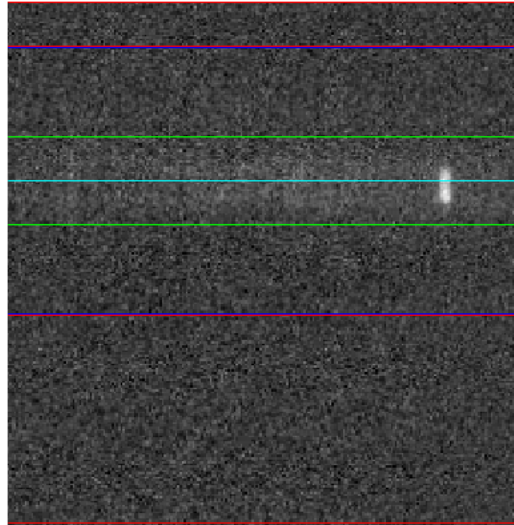

F0432_FO_GRI_05006326_FORG227_TRC_0108.fits

12. Extract Spectra: This module uses the apertures defined in previous modules to extract a 1-d spectrum from each of the input files.

Options: The 'optimal' extraction method is the one we will use; it is also the default option unless the observations are of very extended sources, or if the data were taken in a slit scan mode. The pipeline will also attempt to fix bad pixels; if this causes errors to occur (it doesn't for our example), this option can be turned off.

Output: The output image shows that the bad pixels have been repaired. The cyan line shows the aperture center, the green lines show the edges of the extraction aperture, and the dark blue

lines show the region as the flux goes to zero (PSF radius), and the red lines mark the background regions. The individual spectra are displayed after the next module.


F0432_FO_GRI_05006326_FORG227_SPM_0108.fits

13. Merge Apertures: This module combines the spectra extracted from each aperture into a single spectrum per input file.

Options: In our example, we only have one aperture, but for other observing modes there may be more. Since there is only one aperture, these settings (aside from the 'Save output' option) have no effect on the output data.
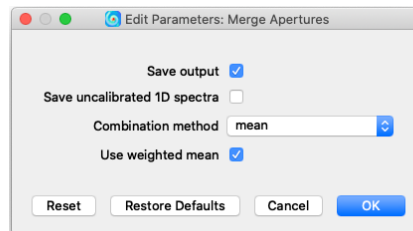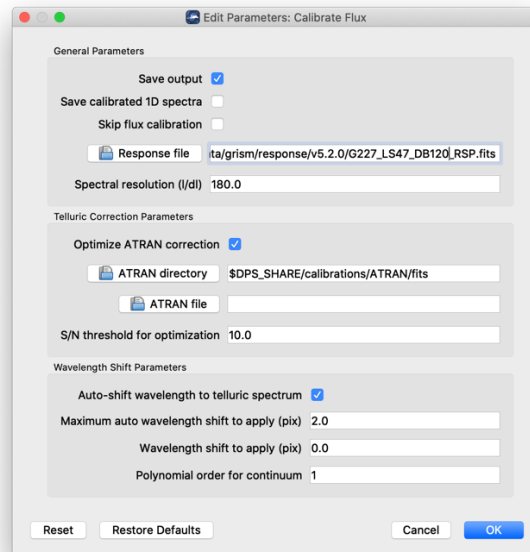

Output: The spectra (one for each input file) are now plotted together in the Eye of SOFIA window, an advanced spectral viewer. Note that the y-axis units are still in Me/s since we have not yet flux calibrated the data, which happens in the next module.
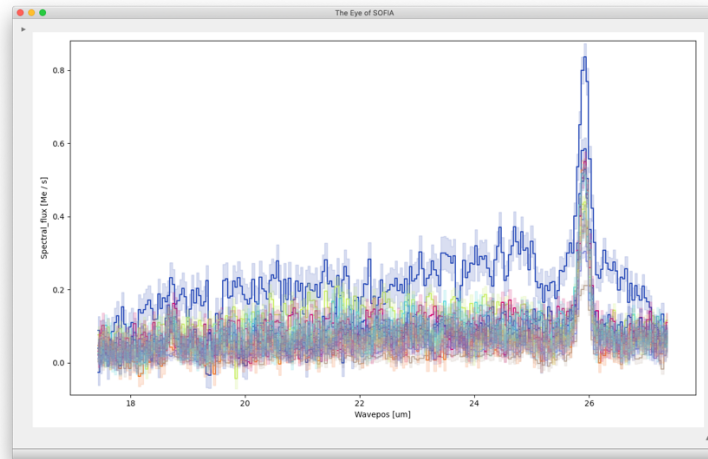
14. Calibrate Flux: In this module, each spectrum is corrected for instrumental response and atmospheric transmission, resulting in a calibrated spectrum in units of Jy per pixel. This module is broken up into three sub-modules as shown below. The telluric correction sub-module corrects for the opacity of the atmosphere using ATRAN models of the atmosphere (courtesy of Steve Lord: https://atran.arc.nasa.gov/cgi-bin/atran/atran.cgi). Due to size constraints in the git repository, the default pipeline installation contains only a single atmospheric model. It's very likely that this file will be insufficient to properly correct for the atmosphere, so you will need to download a library of ATRAN files for a range of flight altitude, telescope elevation, and water vapor. There are two libraries available for download on the https://www.sofia.usra.edu/science/data/data-pipelines page: standard models, which use an assumed water vapor column based on flight altitude and telescope elevation, and specific models, which are calculated for every combination over a range of zenith water vapor values, telescope elevations, and flight altitudes. In this tutorial, we will use the standard models (1) because we have no measurements of water vapor during the observation, and (2) to save download time and disk space. Take the time now to download the standard models, and note their destination directory (if you have trouble finding or downloading the files, please contact the SOFIA Helpdesk at sofia_help@sofia.usra.edu).

Options: In the 'General Parameters' sub-module, the instrumental response file can be specified. The appropriate file should be loaded by default, but a different file can be used if desired. In the 'Telluric Correction Parameters' sub-module, you may choose between ATRAN optimization or a default ATRAN model for telluric correction. ATRAN optimization uses a set of ATRAN files at different Precipitable Water Vapor (PWV) values and appropriate to the observed Zenith Angle and Altitude (the 'specific' files mentioned above) to determine the best PWV to use for telluric correction. In this tutorial, we will de-select the 'Optimize ATRAN correction' option and specify the standard ATRAN files by clicking on the 'ATRAN directory' button and browse to the directory containing the models you downloaded. This sub-module also provides an option to choose a custom user supplied telluric model using the ATRAN file selection button. If a user provided telluric model is used, then all other options are de-selected. The 'Wavelength Shift Parameters' sub-module provides options to apply a wavelength shift to the data to get the

best telluric correction. By default, the pipeline finds the best shift using an optimization routine and tries shift values up to the maximum shift of 2.0 pixels. If the correlation between the transmission and source spectra is not strong, this routine may suggest shifts that are relatively large and not appropriate for the data. Here again the user can use a custom shift using the 'Wavelength shift to apply (pix)' input, which de-selects all the other parameters from this sub-module.


Output: The calibrated spectra (one for each input file, in Jy) are now plotted together in the Eye of SOFIA window.
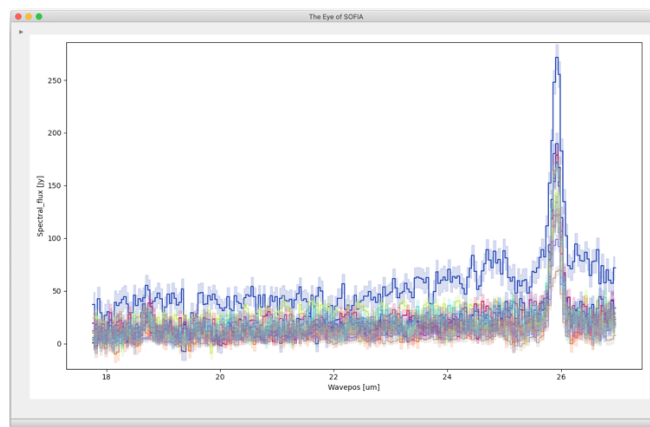

15. Combine Spectra: This module combines the 1-D spectra into a single spectrum and the and 2-D images into a single image.

Options: The options in this module are also broken up into three sub-modules. In the 'General Parameters' sub-module, we can select how the images are registered. The default method of 'Correct to target position' will use the target positions from the headers, if they exist, or else will use the WCS as-is. The combination method determines if the mean (optionally weighted by errors) or median is used. The default is to use the error-weighted mean. In the 1-2D Parameters sub-module, if the 'Robust combination' option is selected, the data are iteratively sigma-clipped before they are combined using the other two settings. The '3D Resample Parameters' are used for slit-scan data when creating 3D data cubes and are not applicable to this example.
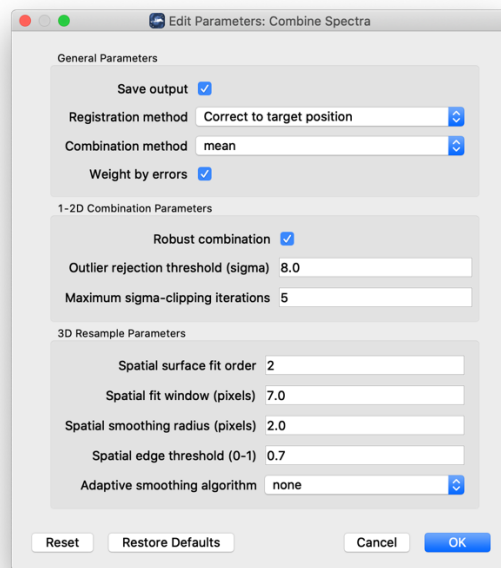
Output: The calibrated, combined spectrum is plotted in the Eye of SOFIA window.
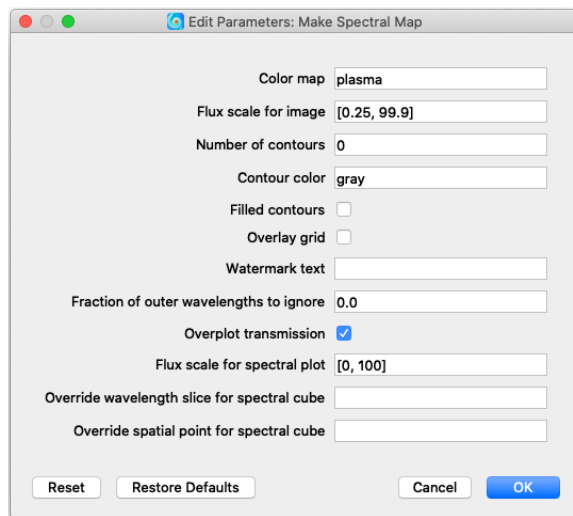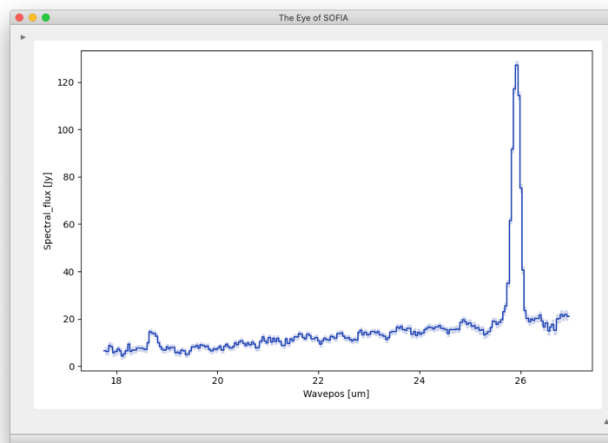
In addition, two final FITS data products are produced, one containing the final 2D image (shown below; the y-axis is the spatial direction, and the x-axis is wavelength) and one containing the final 1D spectrum.
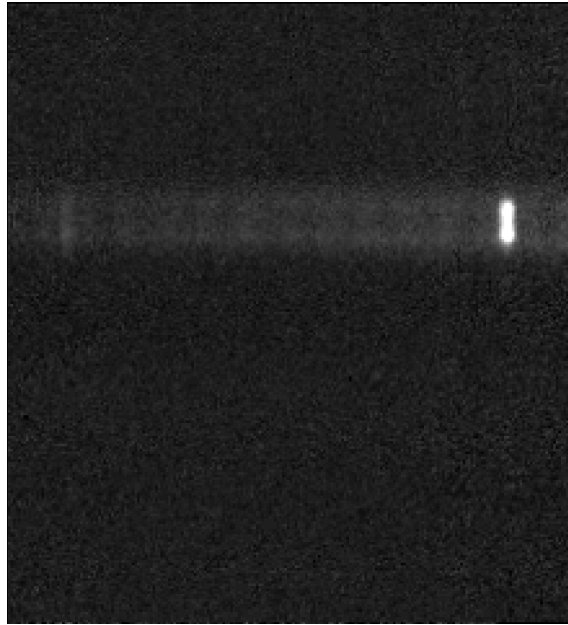
F0432_FO_GRI_05006326_FORG227_COA_0106-0121.fits

16. Make Spectral Map: This last module is used to generate a png file from the final data product.

Options: These options control the look of the final png image.

Output: The final png file contains both the spectrum (overlaid with a plot of the atmospheric transmission) and the 2D image.

Object: NGC 7009, Grism: FOR_G227

Filename: F0432_FO_GRI_05006326_FORG227_COA_0106-0121.fits